

Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis

Sorokin, Dimitry Y.; Makarova, Kira S.; Abbas, Ben; Ferrer, Manuel ; Golyshin, Peter; Galinski, Erwin A.; Ciordia, Sergio; Mena, Maria Carmen; Merkel, Alexander Y.; Wolf, Yuri I.; van Loosdrecht, Marc C.M.; Koonin, Eugene V.

Nature Microbiology

DOI:

[10.1038/nmicrobiol.2017.81](https://doi.org/10.1038/nmicrobiol.2017.81)

Published: 01/05/2017

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Sorokin, D. Y., Makarova, K. S., Abbas, B., Ferrer, M., Golyshin, P., Galinski, E. A., Ciordia, S., Mena, M. C., Merkel, A. Y., Wolf, Y. I., van Loosdrecht, M. C. M., & Koonin, E. V. (2017). Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. *Nature Microbiology*, 2, [17081].
<https://doi.org/10.1038/nmicrobiol.2017.81>

Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis

Dimitry Y. Sorokin^{1,2*}, Kira S. Makarova³, Ben Abbas², Manuel Ferrer⁴, Peter N. Golyshin⁵, Erwin A. Galinski⁶, Sergio Ciordia⁷, María Carmen Mena⁷, Alexander Y. Merkel¹, Yuri I. Wolf³, Mark C.M. van Loosdrecht², Eugene V. Koonin^{3*}

¹*Winogradsky Institute of Microbiology, Centre for Biotechnology, Russian Academy of Sciences, Moscow, Russia;*

²*Department of Biotechnology, Delft University of Technology, Delft, The Netherlands;*

³*National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA;*

⁴*Institute of Catalysis, CSIC, Madrid, Spain;*

⁵*School of Biological Sciences, Bangor University, Gwynedd, UK*

⁶*Institute of Microbiology and Biotechnology, Rheinische Friedrich-Wilhelms University, Bonn, Germany*

⁷*Proteomics Facility, Centro Nacional de Biotecnología, CSIC, Madrid, Spain*

*Corresponding authors:

Dimitry Y. Sorokin: soroc@inmi.ru; d.sorokin@tudelft.nl

Eugene V. Koonin: koonin@ncbi.nlm.nih.gov

25 Methanogenic archaea are major players in the global carbon cycle and in the biotechnology of
26 anaerobic digestion. The phylum *Euryarchaeota* includes diverse groups of methanogens that are
27 interspersed with non-methanogenic lineages. So far methanogens inhabiting hypersaline
28 environments have been identified only within the order *Methanosarcinales*. We report the
29 discovery of a deep phylogenetic lineage of extremophilic methanogens in hypersaline lakes, and
30 present analysis of two nearly complete genomes from this group. Within the phylum
31 *Euryarchaeota*, these isolates form a separate, class-level lineage "Methanonatronarchaeia" that
32 is most closely related to the class *Halobacteria*. Similar to the *Halobacteria*,
33 "Methanonatronarchaeia" are extremely halophilic and do not accumulate organic
34 osmoprotectants. The high intracellular concentration of potassium implies that
35 "Methanonatronarchaeia" employ the "salt-in" osmoprotection strategy. These methanogens
36 are heterotrophic methyl-reducers that utilize C₁-methylated compounds as electron acceptors
37 and formate or hydrogen as electron donors. The genomes contain an incomplete and
38 apparently inactivated set of genes encoding the upper branch of methyl group oxidation to CO₂
39 as well as membrane-bound heterosulfide reductase and cytochromes. These features
40 differentiates "Methanonatronarchaeia" from all known methyl-reducing methanogens. The
41 discovery of extremely halophilic, methyl-reducing methanogens related to haloarchaea provides
42 insight light into the origin of methanogenesis and shows that the strategies employed by
43 methanogens to thrive in salt-saturating conditions are not limited to the classical
44 methylotrophic pathway.
45

Introduction

Methanogenesis is one of the key terminal anaerobic processes of the biogeochemical carbon cycle both in natural ecosystems and in industrial biogas production plants^{1,2}. Biomethane is a major contributor to global warming³. Methanogens comprise four classes, "*Methanomicrobia*", *Methanobacteria*, *Methanopyri* and *Methanococci*, and part of the class *Thermoplasmata*, within the archaeal phylum *Euryarchaeota*^{4,7}. The recent metagenomic discovery of putative methyl-reducing methanogens in the Candidate phyla "*Bathyarchaeota*"⁸ and "*Verstaraetearchaeota*"⁹ indicates that methanogenesis might not be limited to *Euryarchaeota*.

Three major pathways of methanogenesis are known^{1,2}: hydrogenotrophic (H_2 , formate and CO_2 /bicarbonate as electron acceptor), methylotrophic (dismutation of C_1 methylated compounds to methane and CO_2) and acetoclastic (disruption of acetate into methane and CO_2). In the hydrogenotrophic pathway, methane is produced by sequential 6-step reduction of CO_2 . In the methylotrophic pathway, methylated C_1 compounds, including methanol, methylamines and methylsulfides, are first activated by specific methyltransferases. Next, one out of four methyl groups is oxidized through the same reactions as in the hydrogenotrophic pathway occurring in reverse, and the remaining three groups are reduced to methane. In the acetoclastic pathway, methane is produced from the methyl group after activation of acetate. The only enzyme that is uniquely present in all three types of methanogens is methyl-CoM reductase, a Ni-corrinoid protein catalyzing the last step of methyl group reduction to methane¹⁰⁻¹².

The recent discovery of methanogens among *Thermoplasmata*^{5,13-15} drew attention to the fourth, methyl-reducing, pathway, previously characterized in *Methanosphaera stadtmanae* (*Methanobacteria*) and *Methanomicrococcus blatticola* ("*Methanomicrobia*")¹⁶⁻²⁰. In this pathway, C_1 methylated compounds are used only as electron acceptors, whereas H_2 serves as electron donor. In the few known representatives, the genes for methyl group oxidation to CO_2 are either present but inactive (*Methanosphaera*)¹⁶ or completely lost (*Thermoplasmata* methanogens)⁶⁻⁷. Recent metagenomic studies have uncovered three additional, deep lineages of potential methyl-reducing methanogens, namely, Candidate class "*Methanofastidiosa*" within *Euryarchaeota*²¹ and Candidate phyla

"Bathyarchaeota" and "Verstraetearchaeota"^{8,9}, supporting the earlier hypothesis that this is an independently evolved, ancient pathway²².

The classical methylotrophic pathway of methanogenesis that has been characterized in moderately halophilic members of *Methanosarcinales*²³, apparently dominates in hypersaline conditions²³⁻²⁵. In contrast to the extremely halophilic haloarchaea, these microbes only tolerate saturated salt conditions but optimally grow at moderate salinity (below 2-3 M Na⁺) using organic compounds for osmotic balance ("salt-out" strategy)^{26,27}.

Our recent study of methanogenesis in hypersaline soda lakes identified methylotrophic methanogenesis as the most active pathway. In addition, culture-independent analysis of the *mcrA* gene, a unique marker of methanogens, identified a deep lineage that is only distantly related to other methanogens²⁸. We observed no growth of these organisms upon addition of substrates for the classical methanogenic pathways and concluded that they required distinct growth conditions. Here we identify such conditions and describe the discovery and physiological, genomic and phylogenetic features of a previously overlooked group of extremely halophilic, methyl-reducing methanogens.

Discovery of an unknown deep lineage of extremely halophilic methanogens in hypersaline lakes

Sediment stimulation experiments

Two deep-branching *mcrA* sequences have been previously detected in sediments from hypersaline soda lakes in south-eastern Siberia²⁸. Attempts to stimulate the activity of these uncharacterized, dormant methanogens by variation of conditions (temperature, pH and salinity) and substrates elicited a positive response at extreme salinity (4 M Na⁺), pH (9.5-10), elevated temperature (above 48-55°C) and in the presence of methylotrophic substrates together with formate or H₂ (the combination used in the methyl-reducing pathway). The typical response involved a pronounced increase in methane production upon combining methyl compounds with formate or H₂ (less active) compared to single substrates (**Supplementary Figure 1 a**). The *mcrA* profiling of such incubations revealed two distinct clusters closely related to the previously detected deep methanogenic lineage²⁸ (**Supplementary Figure 2**).

The same approach was used with sediment slurries from hypersaline lakes with neutral pH (with no previous evidence of the presence of methyl-reducing methanogens). In this case, enhanced methane production under methyl-reducing conditions (MeOH/trimethylamine + formate) was also observed at elevated temperatures (**Supplementary Figure 1 b,c**). The *mcrA* profiles indicated that typical halophilic methylotrophic methanogens (*Methanohalophilus* and *Methanohalobium*) were outcompeted at high temperature (50-60°C) by unknown, extremely halophilic methyl-reducers which formed a sister clade to the sequences from methyl-reducing incubations of soda lakes sediments in the *mcrA* tree (**Supplementary Figure 2**).

Cultivation of the extremely halophilic methyl-reducing methanogens

The active sediment incubations from hypersaline lakes (**Supplementary Table 1**) were used as an enriched source to obtain the methyl-reducing methanogens in laboratory culture using synthetic media with 2-4 M Na⁺, pH 7 (for salt lakes) or 9.5-10 (for soda lakes), supplemented with MeOH/formate or trimethylamine (TMA)/formate and incubated at 48-60°C. Methane formation was observed only at extreme salinity, close to saturation (4 M total Na⁺), but ceased after the original sediment inoculum was diluted by 2-3 consecutive 1:100 transfers. Addition of colloidal FeS_xnH₂O (soda lakes) or sterilized sediments (salt lakes), combined with filtration through 0.45 µm filters and antibiotic treatment, yielded a pure culture from Siberian soda lakes (strain AMET1 [Alkaliphilic Methylotrophic Thermophilic]), and 10 additional pure AMET cultures from hypersaline alkaline lakes in various geographic locations. A similar approach resulted in three highly enriched cultures at neutral pH from salt lakes (HMET [Halophilic Methylotrophic Thermophilic] cultures) (**Supplementary Table 2**). Phylogenetic analysis of the marker genes showed that AMET and HMET formed two potential genus-level groups that shared 90% 16S rRNA gene sequence identity.

Microbiological characteristics of the methyl-reducing methanogens

Cell morphology and composition

Both AMET and HMET possess small coccoid cells that are motile, in the case of AMET, and lack F₄₂₀ autofluorescence that is typical of most methanogens. A thin, single-layer cell wall was present in

both groups (**Figure 1; Supplementary Figure 3**). At salt concentration below 1.5 M total Na⁺, the cells lost integrity.

The extreme halophily of the discovered methanogens is unprecedented. The salt-tolerant methylotrophs isolated so far from hypersaline habitats, such as *Methanohalobium*, *Methanohalophilus* and *Methanosalsum*, all accumulate organic osmolytes ("salt-out" osmoprotection). In contrast, no recognizable organic osmolytes were detected in AMET1 cells that, instead, accumulated high intracellular concentrations of potassium [5.5 μmol/g protein or 2.2 M, assuming the cell density of 1.2 mg/ml for haloarchaea²⁹ and the measured protein content of 30%]. This concentration is twofold lower than that normally observed inside the cells of haloarchaea (12- 13 μmol/g protein) but close to that of *Halanaerobium* (6.3 μmol/g protein), both of which have been shown to employ the "salt-in" osmoprotection strategy^{30,31}. Furthermore, half of the sodium in the medium is present in the form of carbonates, which possess exactly twofold less osmotic activity than NaCl, resulting in decreased total osmotic pressure, and accordingly, a lower intracellular concentration of osmolytes in extreme natronophiles³². This finding suggests that the extremely halophilic methyl-reducing methanogens rely on potassium as the major osmolyte.

The AMET cell pellets were pinkish in color, suggestive of the presence of cytochromes which was confirmed by difference spectra of a cell-free extract from AMET1 that showed peaks characteristic of *b*-type cytochromes (**Supplementary Figure 4 a**). Given that the cytochrome-containing methanogens of the order *Methanosarcinales* also synthesize the electron-transferring quinone analogue methanophenazine³³, we attempted to detect this compound in AMET1. Indeed, two yellow-colored autofluorescent hydrophobic fractions were recovered from the AMET1 cells, with main masses of 562 and 580 Da, which behaved similar to methanophenazine from *Methanosarcina* (mass 532 Da) upon chemical ionization (sequential cleavage of the 68 Da mass isoprene unit) (**Supplementary Figure 4 b**).

Growth physiology

Both AMET and HMET are methyl-reducing heterotrophic methanogens utilizing C₁-methyl compounds as *e*-acceptor, formate or H₂ as *e*-donor, and yeast extract or acetate as the C-source.

Growth of both groups of organisms was stimulated by addition of external CoM (up to 0.1 mM). Despite the general metabolic similarity, the AMET cultures grew and survived long storage much better than the HMET cultures. The AMET cultures grew best with MeOH as acceptor and formate as donor (**Figure 2 a**). Apart from MeOH, slower growth was also observed with methylamines and dimethylsulfide (**Figure 2 b**). In sharp contrast to the known methyl-reducing methanogens, H₂ was less effective as the electron donor.

Both groups grew optimally around 50°C, with the upper limit at 60°C (**Figure 2 c**, **Supplementary Figure 5**). The AMET isolates were obligate alkaliphiles, with optimum growth at pH 9.5-9.8 (**Figure 2 d**), whereas the HMET cultures had an optimum at pH 6.8-7. The organisms of both groups showed the fastest growth and the highest activity at salt-saturating conditions, and thus qualified as extreme halo(natrono)philes (**Figure 2 e,f**).

Effect of iron sulfides on growth and activity of AMET1

Apart from hydrotroilite (FeS_xnH₂O), AMET1 also grew, albeit less actively, in the presence of crystalline FeS, and yet less actively, with pyrite (FeS₂). No other forms of reduced iron minerals tested (olivine, FeCO₃, magnetite, ferrotine (FeS_n) or various iron(II) silicates could replace FeS. Furthermore, methanogenic activity of resting cells depleted for FeS showed dependence on FeS addition (**Figure 3**). No methane was formed in the absence of either methyl acceptors or formate/H₂, suggesting that Fe²⁺ likely served as a catalyst or regulator rather than a direct *e*-donor. The specific cause(s) of the dependence of AMET growth on iron (II) sulfides remains to be identified.

Comparative genomic analysis

General genome characteristics

The general genome characteristics of AMET1 and HMET1 are given in **Table 1**. Based on analysis of 218 core arCOGs³⁴, both genomes are nearly complete, with two genes missing from this list in AMET1 and three in HMET1. Two of these genes are missing in both genomes (prefoldin paralog GIM5 and deoxyhypusine synthase DYS1), suggesting that they were lost in the common ancestor (**Supplementary Table 3**). The presence of tRNAs for all amino acids is another indication

of genome completeness. The high coverage of the AMET1 and HMET1 genomes by arCOGs implies that the unique phenotype of these organisms is supported largely by the already well-sampled part of the archaeal gene pool.

Phylogenetic analysis and taxonomy

A concatenated alignment of the 56 ribosomal proteins that are universally conserved in complete archaeal genomes³⁵ including AMET1 and HMET1 was used for maximum likelihood tree reconstruction (**Figure 4 a, Supplementary Table 3, Supplementary Data 1**). Both AMET1 and HMET1 belong to a distinct clade, a sister taxon to the class *Halobacteria*, with 100% bootstrap support (**Figure 4 a**). The 16S rRNA gene tree suggests that both organisms belong to the uncultured SA1 group that was first identified in the brine-seawater interface of the Shaban Deep in the Red Sea³⁶ and subsequently in other hypersaline habitats³⁷ (**Supplementary Figure 6**). According to the rRNA phylogeny, the group that includes AMET1 and HMET1 is well separated from the other classes in the phylum Euryarchaeota, both methanogenic and non-methanogenic. The 16S rRNA sequences of these organisms are equally distant from all classes in *Euryarchaeota* and fall within the range of recently recommended values (80-86%) for the class level classification³⁸. Together, these findings appear to justify classification of the SA1 group, including the AMET and HMET lineages, as a separate euryarchaeal class "**Methanonatronarchaeia**". This class would be represented by two distinct genera and species "**Methanonatronarchaeum thermophilum**" (AMET) and '**Candidatus Methanohalarchaeum thermophilum**' (HMET).

Comparative genomic analysis and reconstruction of main evolutionary events

Using arCOG assignments and the results of previous phylogenomic analysis³⁹, we reconstructed the major evolutionary events in the history of AMET1, HMET1 and *Halobacteria* (**Figure 4 a and Supplementary Table 4**). This reconstruction indicates that evolution of the HMET1-AMET1 lineage was dominated by gene loss, whereas *Halobacteria* acquired most of their gene complement after the divergence from "Methanonatronarchaeia". As shown previously, the

common ancestor of *Methanomicrobia* and *Halobacteria* was a methanogen³⁹. The key genes coding for components of the protein complexes involved in the classical methanogenesis pathways, such as tetrahydromethanopterin S-methyltransferase (Mtr), F₄₂₀-reducing hydrogenase and Ftr, appear to have been lost along the branch leading to the common ancestor of *Halobacteria* and "Methanonatronarchaeia". After the divergence, *Halobacteria* continued to lose all other genes involved in methanogenesis and acquire genes for aerobic and mostly heterotrophic pathways, whereas "Methanonatronarchaeia" retained most pathways for anaerobic metabolism, while rewiring the methanogenic pathways for the mixotrophic lifestyle (**Figure 5 a**). As in other cases, genome reduction in "Methanonatronarchaeia" affected RNA modification, DNA repair and stress response systems as well as surface protein structures³⁹. The subsequent gene loss occurred differentially in the two groups of "Methanonatronarchaeia", suggesting adaptation to different ecological niches. The HMET group lost chemotaxis and motility genes and shows signs of adaptation to heterotrophy, whereas AMET retains the ability to synthesize most cellular building blocks at the expense of transporter loss. The AMET strains are motile but lost attachment pili, which are present in the vast majority of the species of the *Halobacteria-Methanomicrobia* clade⁴⁰, and many glycosyltransferases, suggesting simplification of the surface protein structures. The presence of two complete CRISPR-Cas systems in HMET1 compared to none in AMET1, along with the large excess of genes implicated in anti-parasite defense and transposons in HMET1 (**Figure 5 c** and **Table 1**), further emphasize the lifestyle differences indicating that HMET1 is subject to a much stronger pressure from mobile elements than AMET1.

Central metabolism reconstruction

In agreement with the experimental results, genome analysis allowed us to identify the genes of AMET1 and HMET1 that are implicated in energy flow and key reactions of biomass production, which appear to be simple and straightforward (**Figure 6**). The main path starts with utilization of C₁ methyl-containing compounds for methane production by CoM methyltransferases and methyl-CoM reductase complexes, respectively. Similar to the methyl-reducing *Methanomasillicoccales*⁶, the genomes of AMET1 and, especially, HMET1 contains multiple operons encoding diverse

methyltransferases (**Supplementary Figure 7**). Methyl-reduction is coupled with ATP generation and involves five membrane-associated complexes, namely, formate dehydrogenase, membrane-bound heterodisulfide reductase HdrED, Ni,Fe hydrogenase I, multisubunit Na⁺/H⁺ antiporter and H⁺ - transporting ATP synthase. The recently characterized complete biosynthetic pathway⁴¹ for coenzyme F₄₃₀ is present in both genomes. In addition, membrane *b*-type cytochromes and methanophenazine-like compounds are implicated in electron transport.

Pyruvate, the key entry point for biomass production, is generated through acetate incorporation by acetyl-CoA synthetase (**Figure 6**). In a sharp contrast to most methanogens, both genomes lack genes for tetrahydromethanopterin S-methyltransferase Mtr complex and formylmethanofuran dehydrogenase Fwd complex, leaving all intermediate reactions, for which the genes are present, unconnected to other pathways (**Figure 6**). All four recently reported deep lineages of euryarchaeal methyl-reducing methanogens (*Methanomasillilicoccales* and '*Candidatus* Methanofastidiosa') and those from the TACK superphylum ('*Candidatus* Bathyarchaeota' and '*Candidatus* Verstraetearchaeota')^{8,9} lack the Mtr and Fwd complexes as well, but they also lack all the genes involved in intermediate reactions. It is extremely unlikely that genes for all Mtr and Fwd complex subunits are present in both AMET1 and HMET1 but were missed by sequencing. Thus, these organisms might possess still unknown pathways to connect the intermediate reactions to the rest of the metabolic network.

In addition to the main biosynthetic pathway, AMET1 and HMET1 possess genes for three key reactions of anaplerotic CO₂ fixation, namely, malic enzyme, phosphoenolpyruvate carboxylase and carbamoylphosphate synthase. Furthermore, complete gene sets for CO₂ fixation pathway through archaeal RUBISCO are present in both genomes (**Figure 6**)⁴². The great majority of the genes involved in the key biosynthetic pathways for amino acids, nucleotides, cofactors and lipids also were identified in both genomes and found to be highly expressed in proteomic analysis, as revealed by estimating the absolute protein amount based on the exponentially modified protein abundance index (emPAI) (**Supplementary Table 6 and 7**). Interestingly, emPAI-based abundances follow an exponential distribution in which 4 proteins involved in methanogenesis are among the 10 most highly expressed proteins.

HMET1 seems to be more metabolically versatile compared with AMET1, especially with respect to methanogenesis as well as amino acid and sugar metabolism (**Figure 5 b** and **5 c**). However, unlike AMET1, HMET1 lacks several genes for cofactor biosynthesis, such as quinolinate synthase NadA and nicotinate-nucleotide pyrophosphorylase NadC, both involved in NAD biosynthesis; uroporphyrinogen-III decarboxylase HemE and protoporphyrinogen IX oxidase HemG involved in heme biosynthesis, sulfopyruvate decarboxylase involved in CoM biosynthesis and *cofCED* genes for the coenzyme F₄₂₀ biosynthesis enzyme complex. This shortage of biosynthetic enzymes is consistent with experimental observations on poorer growth and survival of HMET in culture compared to AMET.

Adaptation to extreme salinity

Given that acidification of proteins is a common feature of the "salt-in" osmotic strategy, we estimated isoelectric points for the proteomes of a large representative set of halophilic and non-halophilic archaea and bacteria, and compared the distributions as described under Materials and Methods (**Supplementary Table 5**). The distributions of isoelectric points in the AMET1 and HMET1 proteomes are similar to those of moderately halophilic archaea and bacteria, with the notable exception of their closest relatives, the extremely halophilic *Halobacteria*, which form a distinct cloud of extremely acidic proteomes (**Figure 5 d**). This separation indicates that the proteome acidity of *Halobacteria* dramatically changed after the divergence from "Methanonatronarchaea" that appear to be an evolutionary intermediate on the path from methanogens to extreme halophiles. In agreement with the "salt-in" osmoprotection strategy, AMET1 and HMET1 encode a variety of K⁺ transporters (arCOG01960) but show no enrichment of transporters for known organic osmolytes, such as glycine, betaine, ectoine, or glycerol, compared with other archaea (**Supplementary Table 3**). On more general grounds, the "salt-out" strategy appears unlikely and perhaps unfeasible for extremely halophilic secondary anaerobes with relatively low energy yield. Taken together, these considerations suggest that the adaptation of "Methanonatronarchaea" to the extreme salinity relies on the "salt-in" strategy. Whether these organisms possess additional mechanisms for cation-binding to compensate

for the relatively low proteome acidity, remains to be determined, but it is also possible that the main counter-anion, in this case, is Cl⁻.

Analysis of the AMET1 and HMET1 protein complements revealed a major expansion of the UspA family of stress response proteins with likely chaperonin function that could contribute to the structural stability of intracellular proteins (**Supplementary Figure 8**). Finally, we identified several arCOGs consisting of uncharacterized membrane proteins (eg. arCOG04755, arCOG04622, arCOG04619) that are specifically shared by AMET1, HMET1 and the majority of *Halobacteria* (**Supplementary Table 3**). Some of these proteins contain pleckstrin homology domains, which contribute to the mechanical stability of membranes in eukaryotes⁴³ and might play a similar role in "Methanonatronarchaeia".

Notably, AMET1 protein expression analysis showed that the DNA/RNA-binding protein Alba, an archaeal histone and one of the UspA family proteins were among the ten most abundant proteins (**Supplementary Table 6**). These proteins contribute to RNA, DNA and protein stability and might play important roles in supporting growth under extreme salinity conditions.

Implications for the origin of methanogenesis

In previous phylogenetic analyses of the methyl coenzyme M reductase complex (McrABCD) subunits, the topology of the tree for these proteins generally reproduced the ribosomal protein-based phylogeny^{13,22}. In the present phylogenetic analysis that used different protein sets and methods, AMET1/HMET1, *Methanomasilliicoccales*¹³, ANME1 group⁴⁴ and '*Candidatus* Methanofastidiosa' (WSA2 group)²¹ clustered together with high confidence (**Supplementary Figure 9 and 10, Supplementary Data 2, 3 and 4**). This topology differs from the topology of the ribosomal protein tree (**Figure 4 a**). This discrepancy could result from a combination of multiple horizontal transfers of *mcrABCD* genes, differential gain and loss of paralogs, insufficient sampling of rare lineages, and phylogenetic artefacts caused by variation of evolutionary rates. Indeed, we observed a complex evolutionary history of McrA, including many lineage-specific duplications and losses (**Supplementary Figure 9**).

Reconstruction of evolutionary events and mapping the methanogens onto the archaeal tree suggests that the origin of methanogenesis dates back to the common ancestor of archaea, with multiple, independent losses in various clades (**Figure 5 a**). The loss of the methanogenic pathways often proceeds through intermediate stages as clearly observed both in "Methanonatronarchaeia" and *Methanomasilliicoccales* (**Figure 5 a**). Comparison of the gene sets (arCOGs) enriched in different groups of methanogens (**Supplementary Table 3**) using multidimensional scaling revealed distinct patterns of gene loss in "Methanonatronarchaeia", *Methanomasilliicoccales*, ANME1 and '*Candidatus* Bathyarchaeota', in agreement with the independent gene loss scenario (**Figure 5 e**). Notwithstanding these arguments, the possibility that '*Candidatus* Bathyarchaeota' and ANME1 acquired methanogenesis via HGT cannot be ruled out, relegating its origin to the common ancestor of *Euryarchaeota*. Further sampling of diverse archaeal genomes should resolve this issue.

Conclusions

We discovered an unknown, deep euryarchaeal lineage of moderately thermophilic and extremely halo(natrono)philic methanogens that thrive in hypersaline lakes. This group is not monophyletic with the other methanogens but forms a separate, class-level lineage "Methanonatronarchaeia" that is most closely related to *Halobacteria*. The "Methanonatronarchaeia" possess the methyl-reducing type of methanogenesis, where C₁-methylated compounds serve as acceptor and formate or H₂ are external electron donor, but differ from all other methanogens with this type of metabolism in the electron transport mechanism. In contrast to all previously described halophilic methanogens, "Methanonatronarchaeia" grow optimally in saturated salt brines and probably employ potassium-based osmoprotection, similar to extremely halophilic archaea and *Halanaerobiales*. This discovery is expected to have substantial impact on our understanding of biogeochemistry, ecology and evolution of the globally important microbial methanogenesis.

Methods

Samples

Anaerobic sediments (depth from 5 to 15 cm) and near bottom brines were obtained in hypersaline soda and salt lakes in south-western Siberia (Altai region) and south Russia Volgograd region and Crimea) in July of 2013-

2015. The salt concentration varied from 100 to 400 g/l and the pH from 6.5-8 (salt lakes) to 9.8-10.5 (soda lakes). In addition, sediments from Wadi al Natrun alkaline hypersaline lakes in Egypt (October 2000) and alkaline hypersaline Searles Lake in California (April 2005) were used as inoculum in methanogenic enrichment cultures. The details of the lake properties are given in **Supplementary Table 1**. The methanogenic potential activity measurements followed by the *mcrA* analysis have been performed in 1:1 sediment-brine slurries as described previously²⁸.

Enrichment and cultivation conditions

For soda lakes, the sodium carbonate-based mineral media containing 1-4 M total Na⁺ strongly buffered at pH 10^{28,45} was used for enrichments. For salt lakes, the mineral medium containing 4 M NaCl and 0.1 M KCl buffered with 50 mM K phosphates at pH 6.8 was employed. Both media after sterilization were supplied with 1 ml/l of acidic⁴⁶ and alkaline W/Se⁴⁷ trace metal solutions, 1 ml/l of vitamin mix⁴⁶, 4 mM NH₄Cl, 20 mg/l yeast extract and 0.1 mM filter-sterilized CoM. The media were dispensed in serum bottles with butyl rubber stoppers of various capacity at 50 (H₂) - 80% (formate) filled volumes, made anoxic with 5 cycles of argon flushing-evacuation and finally reduced by the addition of 1 mM Na₂S and 1 drop/100 ml of 10% dithionite in 1 M NaHCO₃. H₂ was added on the top of argon atmosphere at 0.5 bar overpressure, formate and methanol - at 50 mM, methylamines - at 10 mM, methyl- and dimethyl sulfides - at 5 mM. In case of methylamines, ammonium was omitted from the basic medium. The incubation temperature varied from 30 to 65°C. Analyses of growth parameters, pH-salt profiling of growth and activity of washed cells, optical and electron microscopy and chemical analyses were performed as described previously^{28,45}.

Biomass composition

The presence of organic compatible solutes was tested by using HPLC and ¹H-NMR after extraction from dry cells with EtOH and the intracellular potassium concentration was quantified by ICP-MS. The presence of the methanophenazine analogues was analyzed in acetone extract from lyophilized cells, followed by TLC separation, reextraction with MeOH-chloroform mixture and MS-MS spectrometry.

Genome sequencing and assembly

The genomic DNA from pure and highly enriched cultures was obtained by using UltraClean Microbial DNA Extraction Kit (MoBio Laboratories). The genome sequencing, assembly and automatic annotation of a pure culture from soda lakes and of a metagenome from a highly enriched salt lake culture was performed by BaseClear (Leiden, The Netherlands) using a combination of Illumina and PacBio platforms. Kmer tetranucleotide frequency analysis was used to identify contigs that are likely belong to HMET1 (meta)genome. Genome completeness has been estimated as described previously⁴⁸.

Genome annotation and sequence analysis

The final gene call has been combined from results by PROKKA⁴⁹ and GeneMarkS⁵⁰ pipelines. All protein-coding genes were assigned to the current archaeal Clusters of Orthologous Groups (arCOGs) as described previously³⁴. Protein annotations were obtained by a combination of arCOGs and PROKKA annotations and, in case of conflict, the respective protein have been manually reanalyzed using PSI-BLAST⁵¹ and HHpred results⁵² and their annotations were modified if necessary. Other genomes for comparative genome analysis were extracted from GenBank (March 2016), and where necessary, ORFs were predicted using GeneMarkS⁵⁰.

Protein sequences were aligned using MUSCLE⁵³. Alignments for the tree reconstruction have been filtered to obtain informative position as described previously³⁴. Approximate maximum likelihood phylogenetic trees were reconstructed using FastTree⁵⁴ and PHYML⁵⁵ methods. The PHYML program was used for the phylogenetic tree reconstruction from an alignment of 51 concatenated ribosomal proteins (287 species, 8072 positions), with the following parameters: LG matrix, gamma distributed site rates, default frequencies which were determined by PROTTEST program⁵⁶. Support values were estimated using an approximate Bayesian method implemented in PhyML. For *mcrA*, multiple alignment (145 sequences and 553 positions) was used for tree reconstruction using PhyML and PROTTEST as described above.

Two sets of genes reconstructed previously using the program COUNT³⁹, which employs a Markov chain gene birth and death model, for the ancestors of *Halobacteriales* *Halobacteriales*/*Methanomicrobiales* were used to infer gene gains and losses on the branches leading to *Halobacteriales*, and the discovered clade of extremely halophilic methanogens. We considered an arCOG to be present in these two clades when the respective COUNT probability was higher than 50%. Further reconstruction was done using a straightforward parsimony approach as explained in detail in **Supplementary Table 4**.

Isoelectric points (pI) of individual proteins were calculated according to Bjellqvist et al.⁵⁷ using the pK values from the EMBOSS suite⁵⁸. Genome-wide distributions of the protein pI were obtained as the probability density estimates at 100 points in the 2.0 – 14.0 pH range using the Gaussian kernel method⁵⁹. Kullback-Leibler

divergence of the pI distributions for the pair of genomes *A* and *B*, $D_{KL}(A|B)$ was computed for all ordered pairs of the set. The distance between the genomes was estimated as $D(A,B) = D(B,A) = (D_{KL}(A|B) + D_{KL}(B|A))/2$ ⁶⁰. The matrix of genome distances was projected into a two-dimensional space using the Classical Multidimensional Scaling method^{61,62} as implemented in the R package⁶³.

Proteomics

Proteomic analyses were conducted using the soda lake pure culture AMET1 (48°C, MeOH+formate) and the salt lake enrichment HMET1 (37°C, TMA+H₂) (**Supplementary Table 6 and 7**). Cell pellets were dissolved in lysis buffer (8 M urea, 2 M thiourea, 5% CHAPS, 5 mM TCEP-HCl and a protease inhibitors cocktail). Homogenization of the cells was achieved by ultra-sonication for 5 min on ultrasonic bath. After homogenization, the lysed cells were centrifuged at 20,000×g for 10 min at 4 °C, and the supernatant containing the solubilized proteins was used for LC-MS/MS experiment. All samples were precipitated by methanol/chloroform method and re-suspended in a multi-chaotropic sample solution (7 M urea, 2 M thiourea, 100 mM TEAB; pH 7.5). Total protein concentration was determined using Pierce 660 nm protein assay (Thermo). 40 µg of protein from each sample were reduced with 2 µL of 50mM Tris(2-carboxyethyl) phosphine (TCEP, SCIEX), pH 8.0, at 37°C for 60 min and followed by 1 µL of 200mM cysteine-blocking reagent (methyl methanethiosulfonate (MMTS, Pierce) for 10 min at room temperature. Samples were diluted up to 140 µL to reduce urea concentration with 25mM TEAB. Digestions were initiated by adding 2 µg Pierce MS-grade trypsin (Thermo Scientific) to each sample in a ratio 1:20 (w/w), which were then incubated at 37°C overnight on a shaker. Sample digestions were evaporated to dryness in a vacuum concentrator and then desalted onto StageTip C18 Pipette tips (Thermo Scientific) until the mass spectrometric analysis.

A 1 µg aliquot of each sample was subjected to 1D-nano LC ESI-MSMS analysis using a nano liquid chromatography system (Eksigent Technologies nanoLC Ultra 1D plus, AB SCIEX, Foster City, CA) coupled to high speed Triple TOF 5600 mass spectrometer (SCIEX, Foster City, CA) with a Nanospray III source. The analytical column used was a silica-based reversed phase Acquity UPLC M-Class Peptide BEH C18 Column, 75 µm × 150 mm, 1.7 µm particle size and 130 Å pore size (Waters). The trap column was a C18 Acclaim PepMapTM 100 (Thermo Scientific), 100 µm × 2 cm, 5 µm particle diameter, 100 Å pore size, switched on-line with the analytical column. The loading pump delivered a solution of 0.1% formic acid in water at 2 µl/min. The nano-pump provided a flow-rate of 250 nl/min and was operated under gradient elution conditions. Peptides were separated using a 250 minutes gradient ranging from 2% to 90% mobile phase B (mobile phase A: 2% acetonitrile, 0.1% formic acid; mobile phase B: 100% acetonitrile, 0.1% formic acid). Injection volume was 5 µl.

Data acquisition was performed with a TripleTOF 5600 System (SCIEX, Foster City, CA). Data was acquired using an ionspray voltage floating (ISVF) 2300 V, curtain gas (CUR) 35, interface heater temperature (IHT) 150, ion source gas 1 (GS1) 25, declustering potential (DP) 100 V. All data was acquired using information-dependent acquisition (IDA) mode with Analyst TF 1.7 software (SCIEX, Foster City, CA). For IDA parameters, 0.25s MS survey scan in the mass range of 350–1250 Da were followed by 35 MS/MS scans of 100ms in the mass range of 100–1800 (total cycle time: 4 s). Switching criteria were set to ions greater than mass to charge ratio (m/z) 350 and smaller than m/z 1250 with charge state of 2–5 and an abundance threshold of more than 90 counts (cps). Former target ions were excluded for 15s. IDA rolling collision energy (CE) parameters script was used for automatically controlling the CE.

MS and MS/MS data obtained for individual samples were processed using Analyst® TF 1.7 Software (SCIEX). The reconstituted AMET1 and HMET1 chromosome sequence was used to generate the database for protein identification using the Mascot Server v. 2.5.1 (Matrix Science, London, UK). Search parameters were set as follows: carbamidomethyl (C) as fixed modification and acetyl (Protein N-term), Gln to pyro-Glu (N-term Q), Glu to pyro-Glu (N-term E) and methionine oxidation as variable modifications. Peptide mass tolerance was set to 25 ppm and 0.05 Da for fragment masses, also 2 missed cleavages were allowed. The confidence interval for protein identification was set to ≥ 95% (p<0.05) and only peptides with an individual ion score above the 1% False Discovery Rates (FDR) at PSM level were considered correctly identified. False Discovery Rates were manually calculated. The threshold of only one identified peptide per protein identification was used because FDR controlled experiments counter intuitively suffer from the two-peptide rule⁶⁴. To rank the protein abundance in each sample, the Exponentially Modified Protein Abundance Index (emPAI) was used in the present study as a relative quantitation score of the proteins in a complex mixture based on protein coverage by the peptide matches in a database search result⁶². Although the emPAI is not as accurate as quantification using synthesized peptide standards, it is quite useful for obtaining a broad overview of proteome profiles.

Data availability

The final, assembled and annotated genomic sequences of the two isolates of "Methanonatronarchaea" were deposited in GenBank under accession numbers: PRJNA356895 (MAET1)

and PRJNA357090 (HMET1). Other data supporting the findings reported in this article are available from the Supplementary Information tables or from the corresponding authors upon request.

Acknowledgements. DYS was supported by STW (project 12226), Gravitation-SIAM Program (grant 24002002, Dutch Ministry of Education and Science) and by RFBR (grant 16-04-00035). KSM, YIW and EVK are supported by the intramural program of the U.S. Department of Health and Human Services (to the National Library of Medicine). The proteomic analysis was performed in the Proteomics Facility of The Spanish National Center for Biotechnology (CNB-CSIC) that belongs to ProteoRed, PRB2-ISCI, supported by grant PT13/0001. This project has received funding from the European Union's Horizon 2020 research and innovation program [Blue Growth: Unlocking the potential of Seas and Oceans] under grant agreement No [634486]. This work was further funded by grant BIO2014-54494-R from the Spanish Ministry of Economy, Industry and Competitiveness.

Author contributions. D.Y.S. performed the field work, the sediment activity incubations, enrichment and isolation of pure cultures and microbiological investigation of enriched and pure cultures. B.A. and A.Y.M. analyzed the *mcrA* and 16S rRNA genes in sediments and methanogenic cultures. M.F., P.N.G., S.C. and M.C.M. were responsible for the proteomic analysis. E.G. analyzed compatible solutes. K.S.M., Y.I.W. and E.V.K. performed genomic analysis and evolutionary reconstructions. D.Y.S., K.S.M. and E.V.K. wrote the paper. M.C.M.L. oversaw the project and participated in the data interpretation and discussion.

Competing interests. The authors declare no competing financial interests.

493 **References**

- 494 1 Ferry, J. G. & Kastead, K. A. in *Archaea: Molecular and Cellular Biology* (ed R.
495 Cavicchioli) 288-214 (ASM Press, 2007).
- 496 2 Conrad, R. The global methane cycle: recent advances in understanding the microbial
497 processes involved. *Environmental microbiology reports* **1**, 285-292,
498 doi:10.1111/j.1758-2229.2009.00038.x (2009).
- 499 3 Agency, U. S. E. P. (ed US-EPA) (2016).
- 500 4 Garrity, G. M. & Holt, J. G. in *Bergey's Manual of Systematics of Archaea and*
501 *Bacteria* Vol. 1 (John Wiley & Sons, Inc. , 2015).
- 502 5 Iino, T. *et al.* Candidatus Methanogranum caenicola: a novel methanogen from the
503 anaerobic digested sludge, and proposal of Methanomassiliicoccaceae fam. nov. and
504 Methanomassiliicoccales ord. nov., for a methanogenic lineage of the class
505 Thermoplasmata. *Microbes and environments* **28**, 244-250 (2013).
- 506 6 Borrel, G. *et al.* Comparative genomics highlights the unique biology of
507 Methanomassiliicoccales, a Thermoplasmatales-related seventh order of methanogenic
508 archaea that encodes pyrrolysine. *BMC genomics* **15**, 679, doi:10.1186/1471-2164-15-
509 679 (2014).
- 510 7 Lang, K. *et al.* New mode of energy metabolism in the seventh order of methanogens
511 as revealed by comparative genome analysis of "Candidatus methanoplasma
512 termitum". *Applied and environmental microbiology* **81**, 1338-1352,
513 doi:10.1128/AEM.03389-14 (2015).
- 514 8 Evans, P. N. *et al.* Methane metabolism in the archaeal phylum Bathyarchaeota
515 revealed by genome-centric metagenomics. *Science* **350**, 434-438,
516 doi:10.1126/science.aac7745 (2015).
- 517 9 Vanwonterghem, I. *et al.* Methylophilic methanogenesis discovered in the archaeal
518 phylum Verstraetearchaeota. *Nature microbiology* **1**, 16170,
519 doi:10.1038/nmicrobiol.2016.170 (2016).
- 520 10 Hedderich, R. & Whitman, W. B. in *The Prokaryotes – Prokaryotic Physiology and*
521 *Biochemistry* (ed E. Rosenberg) 636-663 (Springer-Verlag, 2013).
- 522 11 Liu, Y. & Whitman, W. B. Metabolic, phylogenetic, and ecological diversity of the
523 methanogenic archaea. *Annals of the New York Academy of Sciences* **1125**, 171-189,
524 doi:10.1196/annals.1419.019 (2008).
- 525 12 Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W. & Hedderich, R. Methanogenic
526 archaea: ecologically relevant differences in energy conservation. *Nature reviews.*
527 *Microbiology* **6**, 579-591, doi:10.1038/nrmicro1931 (2008).
- 528 13 Borrel, G. *et al.* Phylogenomic data support a seventh order of Methylophilic
529 methanogens and provide insights into the evolution of Methanogenesis. *Genome*
530 *biology and evolution* **5**, 1769-1780, doi:10.1093/gbe/evt128 (2013).
- 531 14 Dridi, B., Fardeau, M. L., Ollivier, B., Raoult, D. & Drancourt, M.
532 *Methanomassiliicoccus luminyensis* gen. nov., sp. nov., a methanogenic archaeon
533 isolated from human faeces. *International journal of systematic and evolutionary*
534 *microbiology* **62**, 1902-1907, doi:10.1099/ijls.0.033712-0 (2012).
- 535 15 Paul, K., Nonoh, J. O., Mikulski, L. & Brune, A. "Methanoplasmales,"
536 Thermoplasmatales-related archaea in termite guts and other environments, are the
537 seventh order of methanogens. *Applied and environmental microbiology* **78**, 8245-
538 8253, doi:10.1128/AEM.02193-12 (2012).
- 539 16 Fricke, W. F. *et al.* The genome sequence of *Methanosphaera stadtmanae* reveals why
540 this human intestinal archaeon is restricted to methanol and H₂ for methane formation

541 and ATP synthesis. *Journal of bacteriology* **188**, 642-658, doi:10.1128/JB.188.2.642-
542 658.2006 (2006).

543 17 Miller, T. L. & Wolin, M. J. *Methanosphaera stadtmaniae* gen. nov., sp. nov.: a
544 species that forms methane by reducing methanol with hydrogen. *Archives of*
545 *microbiology* **141**, 116-122 (1985).

546 18 Sprenger, W. W., Hackstein, J. H. & Keltjens, J. T. The energy metabolism of
547 *Methanomicrococcus blatticola*: physiological and biochemical aspects. *Antonie van*
548 *Leeuwenhoek* **87**, 289-299, doi:10.1007/s10482-004-5941-5 (2005).

549 19 Sprenger, W. W., Hackstein, J. H. & Keltjens, J. T. The competitive success of
550 *Methanomicrococcus blatticola*, a dominant methylotrophic methanogen in the
551 cockroach hindgut, is supported by high substrate affinities and favorable
552 thermodynamics. *FEMS microbiology ecology* **60**, 266-275, doi:10.1111/j.1574-
553 6941.2007.00287.x (2007).

554 20 Sprenger, W. W., van Belzen, M. C., Rosenberg, J., Hackstein, J. H. & Keltjens, J. T.
555 *Methanomicrococcus blatticola* gen. nov., sp. nov., a methanol- and methylamine-
556 reducing methanogen from the hindgut of the cockroach *Periplaneta americana*.
557 *International journal of systematic and evolutionary microbiology* **50 Pt 6**, 1989-
558 1999, doi:10.1099/00207713-50-6-1989 (2000).

559 21 Nobu, M. K., Narihiro, T., Kuroda, K., Mei, R. & Liu, W. T. Chasing the elusive
560 Euryarchaeota class WSA2: genomes reveal a uniquely fastidious methyl-reducing
561 methanogen. *The ISME journal* **10**, 2478-2487, doi:10.1038/ismej.2016.33 (2016).

562 22 Borrel, G., Adam, P. S. & Gribaldo, S. Methanogenesis and the Wood-Ljungdahl
563 Pathway: An Ancient, Versatile, and Fragile Association. *Genome biology and*
564 *evolution* **8**, 1706-1711, doi:10.1093/gbe/evw114 (2016).

565 23 McGenity, T. J. in *Handbook of Hydrocarbon and Lipid Microbiology* (ed K.N.
566 Timmis) 665-679 (Springer-Verlag, 2010).

567 24 Kelley, C. A., Poole, J. A., Tazaz, A. M., Chanton, J. P. & Bebout, B. M. Substrate
568 limitation for methanogenesis in hypersaline environments. *Astrobiology* **12**, 89-97,
569 doi:10.1089/ast.2011.0703 (2012).

570 25 Oremland, R. S. & King, G. M. in *Microbial mats. Physiological ecology of benthic*
571 *microbial communities* (eds Y. Cohen & E. Rosenberg) 180-190 (American Society
572 for Microbiology, 1989).

573 26 Martin, D. D., Ciulla, R. A. & Roberts, M. F. Osmoadaptation in archaea. *Applied and*
574 *environmental microbiology* **65**, 1815-1825 (1999).

575 27 Menaia, J. A. G. F. Osmotics of halophilic methanogenic archaeobacteria. *Scholar*
576 *Archive paper* (1992).

577 28 Sorokin, D. Y. *et al.* Methanogenesis at extremely haloalkaline conditions in the soda
578 lakes of Kulunda Steppe (Altai, Russia). *FEMS microbiology ecology* **91**,
579 doi:10.1093/femsec/fiv016 (2015).

580 29 Ginzburg, M., Sachs, L. & Ginzburg, B. Z. Ion metabolism in a Halobacterium. I.
581 Influence of age of culture on intracellular concentrations. *The Journal of general*
582 *physiology* **55**, 187-207 (1970).

583 30 Elevi Bardavid, R. & Oren, A. The amino acid composition of proteins from anaerobic
584 halophilic bacteria of the order Halanaerobiales. *Extremophiles : life under extreme*
585 *conditions* **16**, 567-572, doi:10.1007/s00792-012-0455-y (2012).

586 31 Oren, A. Life at high salt concentrations, intracellular KCl concentrations, and acidic
587 proteomes. *Frontiers in microbiology* **4**, 315, doi:10.3389/fmicb.2013.00315 (2013).

588 32 Sorokin, D. Y., Banciu, H. L. & Muyzer, G. Functional microbiology of soda lakes.
589 *Current opinion in microbiology* **25**, 88-96, doi:10.1016/j.mib.2015.05.004 (2015).

- 590 33 Abken, H. J. *et al.* Isolation and characterization of methanophenazine and function of
591 phenazines in membrane-bound electron transport of *Methanosarcina mazei* Go1.
592 *Journal of bacteriology* **180**, 2027-2032 (1998).
- 593 34 Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Archaeal Clusters of Orthologous Genes
594 (arCOGs): An Update and Application for Analysis of Shared Features between
595 Thermococcales, Methanococcales, and Methanobacteriales. *Life (Basel)* **5**, 818-840,
596 doi:life5010818 [pii] 10.3390/life5010818 (2015).
- 597 35 Yutin, N., Puigbo, P., Koonin, E. V. & Wolf, Y. I. Phylogenomics of prokaryotic
598 ribosomal proteins. *PLoS One* **7**, e36972, doi:10.1371/journal.pone.0036972 PONE-
599 D-11-23203 [pii] (2012).
- 600 36 Eder, W., Schmidt, M., Koch, M., Garbe-Schonberg, D. & Huber, R. Prokaryotic
601 phylogenetic diversity and corresponding geochemical data of the brine-seawater
602 interface of the Shaban Deep, Red Sea. *Environmental microbiology* **4**, 758-763
603 (2002).
- 604 37 Jiang, H. *et al.* Microbial response to salinity change in Lake Chaka, a hypersaline
605 lake on Tibetan plateau. *Environmental microbiology* **9**, 2603-2621,
606 doi:10.1111/j.1462-2920.2007.01377.x (2007).
- 607 38 Yarza, P. *et al.* Uniting the classification of cultured and uncultured bacteria and
608 archaea using 16S rRNA gene sequences. *Nature reviews. Microbiology* **12**, 635-645,
609 doi:10.1038/nrmicro3330 (2014).
- 610 39 Wolf, Y. I., Makarova, K. S., Yutin, N. & Koonin, E. V. Updated clusters of
611 orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of
612 horizontal gene transfer. *Biol Direct* **7**, 46, doi:10.1186/1745-6150-7-46 1745-6150-7-
613 46 [pii] (2012).
- 614 40 Makarova, K. S., Koonin, E. V. & Albers, S. V. Diversity and Evolution of Type IV
615 pili Systems in Archaea. *Frontiers in microbiology* **7**, 667,
616 doi:10.3389/fmicb.2016.00667 (2016).
- 617 41 Zheng, K., Ngo, P. D., Owens, V. L., Yang, X. P. & Mansoorabadi, S. O. The
618 biosynthetic pathway of coenzyme F430 in methanogenic and methanotrophic
619 archaea. *Science* **354**, 339-342, doi:10.1126/science.aag2947 (2016).
- 620 42 Aono, R. *et al.* Enzymatic characterization of AMP phosphorylase and ribose-1,5-
621 bisphosphate isomerase functioning in an archaeal AMP metabolic pathway. *Journal*
622 *of bacteriology* **194**, 6847-6855, doi:10.1128/JB.01335-12 (2012).
- 623 43 Baines, A. J. Evolution of spectrin function in cytoskeletal and membrane networks.
624 *Biochemical Society transactions* **37**, 796-803, doi:10.1042/BST0370796 (2009).
- 625 44 Hallam, S. J., Girguis, P. R., Preston, C. M., Richardson, P. M. & DeLong, E. F.
626 Identification of methyl coenzyme M reductase A (mcrA) genes associated with
627 methane-oxidizing archaea. *Applied and environmental microbiology* **69**, 5483-5491
628 (2003).
- 629 45 Sorokin, D. Y. *et al.* *Methanosalsum natronophilum* sp. nov., and *Methanocalculus*
630 *alkaliphilus* sp. nov., haloalkaliphilic methanogens from hypersaline soda lakes.
631 *International journal of systematic and evolutionary microbiology* **65**, 3739-3745,
632 doi:10.1099/ijsem.0.000488 (2015).
- 633 46 Pfennig, N. & Lippert, K. D. Über das Vitamin B12-Bedürfnis phototropher
634 Schwefelbakterien. *Arch. Mikrobiol.* **55**, 245-256 (1966).
- 635 47 Plugge, C. M. Anoxic media design, preparation, and considerations. *Methods in*
636 *enzymology* **397**, 3-16, doi:10.1016/S0076-6879(05)97001-8 (2005).
- 637 48 Podar, M. *et al.* Insights into archaeal evolution and symbiosis from the genomes of a
638 nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone
639 National Park. *Biology direct* **8**, 9, doi:10.1186/1745-6150-8-9 (2013).

640 49 Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-
641 2069, doi:10.1093/bioinformatics/btu153 (2014).

642 50 Besemer, J., Lomsadze, A. & Borodovsky, M. GeneMarkS: a self-training method for
643 prediction of gene starts in microbial genomes. Implications for finding sequence
644 motifs in regulatory regions. *Nucleic Acids Res* **29**, 2607-2618 (2001).

645 51 Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein
646 database search programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).

647 52 Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein
648 homology detection and structure prediction. *Nucleic Acids Res* **33**, W244-248,
649 doi:10.1093/nar/gki408 (2005).

650 53 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high
651 throughput. *Nucleic Acids Res* **32**, 1792-1797 (2004).

652 54 Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-
653 likelihood trees for large alignments. *PLoS One* **5**, e9490,
654 doi:10.1371/journal.pone.0009490 (2010).

655 55 Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood
656 phylogenies: assessing the performance of PhyML 3.0. *Systematic biology* **59**, 307-
657 321, doi:10.1093/sysbio/syq010 (2010).

658 56 Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProtTest 3: fast selection of best-
659 fit models of protein evolution. *Bioinformatics* **27**, 1164-1165,
660 doi:10.1093/bioinformatics/btr088 (2011).

661 57 Bjellqvist, B. *et al.* The focusing positions of polypeptides in immobilized pH
662 gradients can be predicted from their amino acid sequences. *Electrophoresis* **14**, 1023-
663 1031 (1993).

664 58 Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open
665 Software Suite. *Trends in genetics : TIG* **16**, 276-277 (2000).

666 59 Parzen, E. On Estimation of a Probability Density Function and Mode. *Ann. Math. Statist.* **33**, 1065-1076 (1962).

667 60 Kullback, S. & Leibler, R. A. On information and sufficiency. *Ann. Math. Stat.* **22**, 79-
668 86 (1951).

669 61 Gower, J. C. Some distance properties of latent root and vector methods used in
670 multivariate analysis. *Biometrika* **53** (1966).

671 62 Torgeson, W. S. *Theory and Methods of Scaling*. (Wiley 1958).

672 63 Team, R. C. (R Foundation for Statistical Computing, Vienna, Austria, 2013).

673 64 Gupta, N. & Pevzner, P. A. False discovery rates of protein identifications: a strike
674 against the two-peptide rule. *Journal of proteome research* **8**, 4173-4181,
675 doi:10.1021/pr9004794 (2009).

676
677
678
679

Figures legends

Fig. 1 Cell morphology of the methyl-reducing methanogens from hypersaline soda (strain AMET1, **a-d**) and salt (strain HMET1, **e-f**) lakes. **a** - phase contrast image; **b** and **e** - total electron microscopy images; **c-d** and **f** - electron microscopy images of thin sectioned cells. N - nucleoid; PHA? - a possible PHA storage granule; CPM - cytoplasmic membrane; ICPM - cell membrane invaginations; CW - cell wall. The light microscopy images are typical of 5-7 samples from 2 replicate cultures; the electron microscopy images are typical from technical replicates from the same culture (n=5 to n=10).

Fig. 2 Growth and activity of methyl-reducing methanogens from hypersaline soda lakes. **a** - growth dynamics of strain AMET1 with MeOH+formate at 4 M total Na^+ , pH 9.5 and 50°C ($Y_{\text{max}}=1.5$ mg protein/mM MeOH; $\mu_{\text{max}}=0.012\text{-}0.015$ h⁻¹). **b** - methanogenic activity of washed cells of strain AMET1 grown with MeOH+formate (at 4 M total Na^+ , pH 9.5 and 48°C) with various methylated *e*-acceptors. **c** - influence of temperature on growth and activity of washed cells of various AMET strains at 4 M total Na^+ , pH 9.5 with MeOH+formate as substrate. **d** - influence of pH at 4 M Na^+ on growth and activity of washed cells of strain AMET1 at 48°C with MeOH+formate as substrate. **e** - influence of salinity at pH 9.5 on growth and activity of washed cells of strain AMET1 with MeOH+formate as substrate. In all experiments, 100 μM hydrotroilite ($\text{FeS} \times n\text{H}_2\text{O}$) was added. Neither growth nor activity were observed with a single substrate (i.e. methylated compounds, H_2 or formate alone). VCH_4 is a rate of methane formation, normalized either per culture volume in growth experiments or per biomass in cell suspension experiments. The error bars in **2a**, **2b** and **2d** are SD of biological replicates (n=2). The plots in **2c** and **2e** represent results of a single sample analysis.

Fig. 3 Effect of hydrotroilite ($\text{FeS} \times n\text{H}_2\text{O}$) on growth and methanogenic activity of washed and exhausted cells of the AMET1 strain

Growth and incubation conditions: 4 M total Na^+ , pH 9.8, 48°C. Substrate: 50 mM CH_3OH +50 mM formate. The culture was grown in the presence of sterile sand. The FeS -exhausted, washed cells were obtained by prolonged incubation with substrates without addition of FeS followed by washing and resuspension in a fresh buffer. The error bars are SD of biological replicates (n=3).

Fig. 4 Phylogenetic analysis of "Methanonatronarchaeia" (AMET1 and HMET1)

The tree represents a phylogeny of archaea based on an alignment of concatenated ribosomal proteins. Methanogen clades are shown in blue and *Halobacteriales* in orange. The inferred methanogenic branches are highlighted in blue, the inferred loss of methanogenesis is indicated by dashed red branches. The arrow indicates the likely archaeal root. All branches are bootstrap-supported at 100% level. The original tree is available in **Supplementary Data 1**.

Fig. 5 Comparative genomic analysis and reconstruction of gene losses and gains

a. Reconstruction of gene loss and gain along in "Methanonatronarchaeia" (AMET1 and HMET1) and *Halobacteriales*. Light blue: arCOG complement; green: gains; dark blue: losses.

b. arCOG composition of AMET1 and HMET1.

c. Distribution of the differences in the arCOG composition of AMET1 and HMET1 by functional categories. For each category, the number of arCOGs unique to AMET1 and HMET1 is indicated. The functional classification of the COGs is described at <ftp://ftp.ncbi.nih.gov/pub/wolf/COGs/arCOG/funcclass.tab>

d. Multidimensional scaling analysis of isoelectric point distributions. Orange: *Halobacteria*; green: halophilic archaea and bacteria; blue: other archaea and bacteria.

e. Multidimensional scaling analysis of genes enriched in methanogens.

Fig. 6 Reconstruction of the central metabolic pathways shared by "Methanonatronarchaeia"

The main methyl-reducing pathway is shown by thick magenta arrows. Metabolically fixed low molecular weight compounds are shown in green. Either gene name or respective arCOG number is shown for each reaction and shown in red (details are available at **Supplementary Table 3**). Final biosynthetic products are shown as follows: light blue for amino acids, pale yellow for nucleotides, brown for lipid components, pink for cofactors. Abbreviations: MF, methanofuran; H4MPT, tetrahydromethanopterin; CoM, coenzyme M, CoB – coenzyme B, CoA – coenzyme A.

Table 1. Summary statistics for the AMET1 and HMET1 genomes.

	AMET1	HMET1
Number of contigs	8	4
Total length (base pairs)	1513137	2141311
Number of coding sequences	1514	2168
GC content	38%	35%
rRNAs	5S, 16S, 23S	5S, 16S, 23S
tRNAs (for different amino acids including pyrrolysine)	31 (21)	37 (21)
Proteins assigned to arCOGs	88%	79%
Completeness based on archaeal core arCOGs	99% [#]	99% [#]
CRISPR arrays	0	4
CRISPR-cas system subtypes	-	I-D, III-B
Transposon-related genes	4 [*]	121 [*]
Integrated elements (His2-like viruses)	3	2

*- Some are probably pseudogenes; # - based on 218 core arCOGs

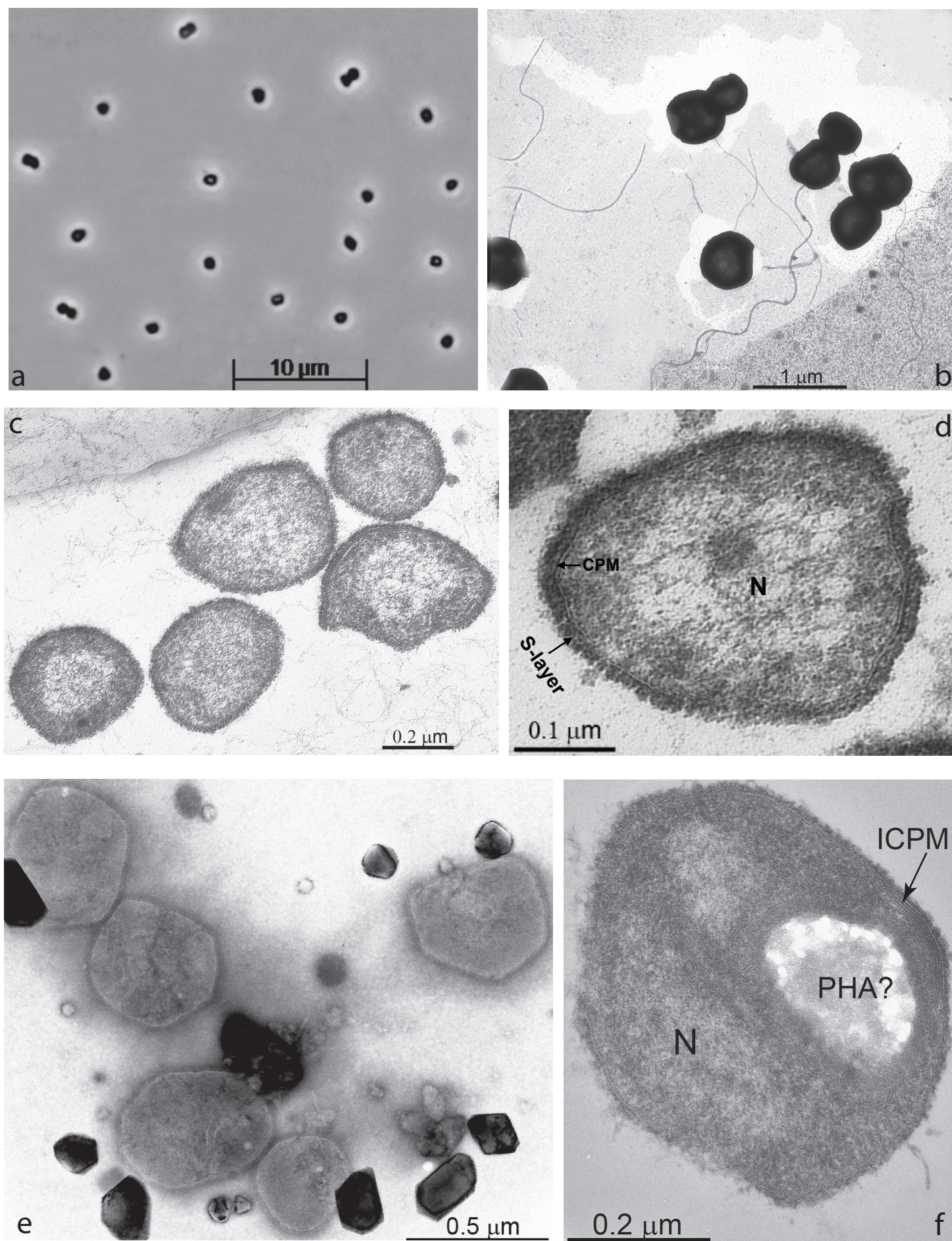


Fig. 1

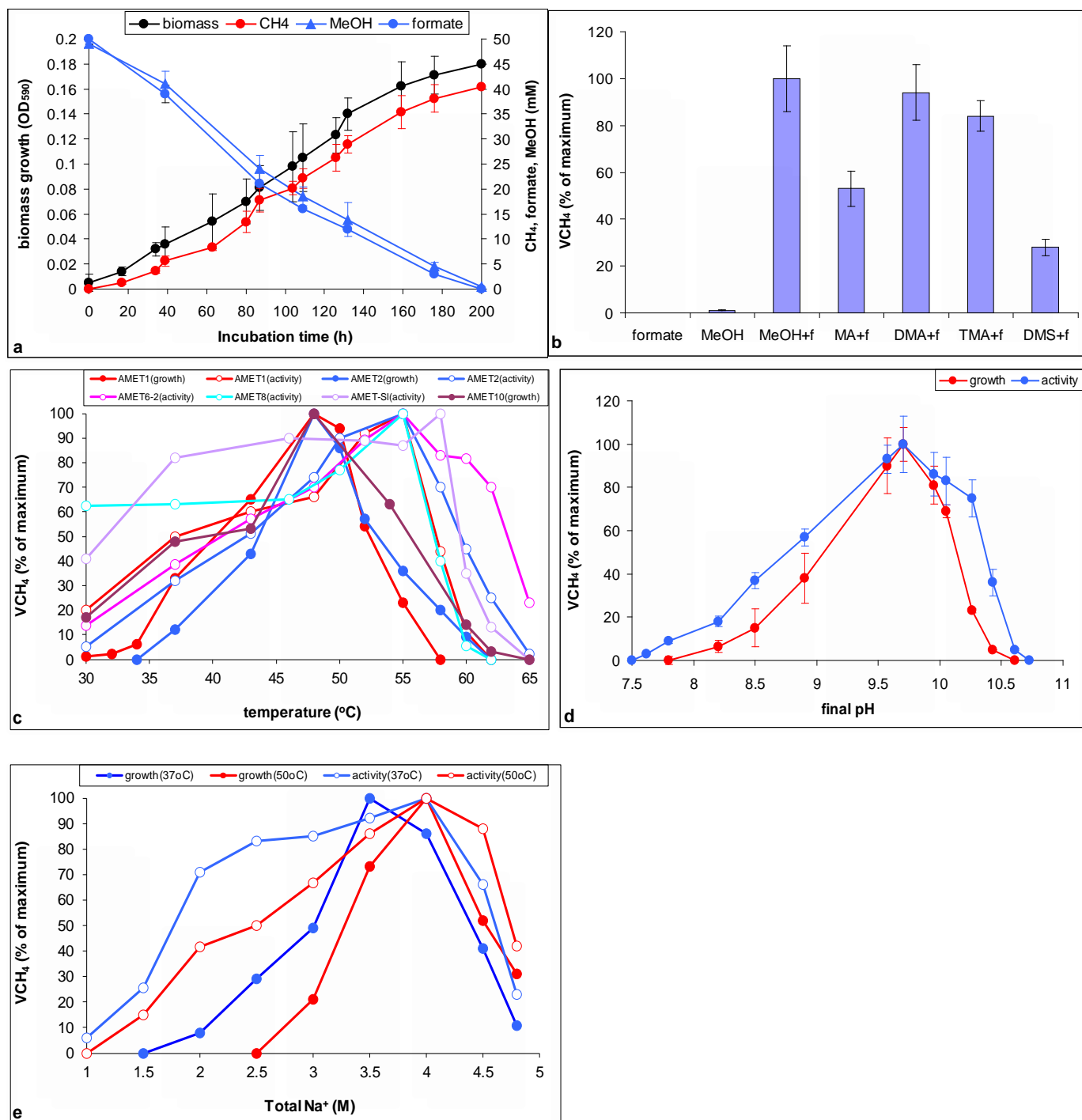


Fig.2

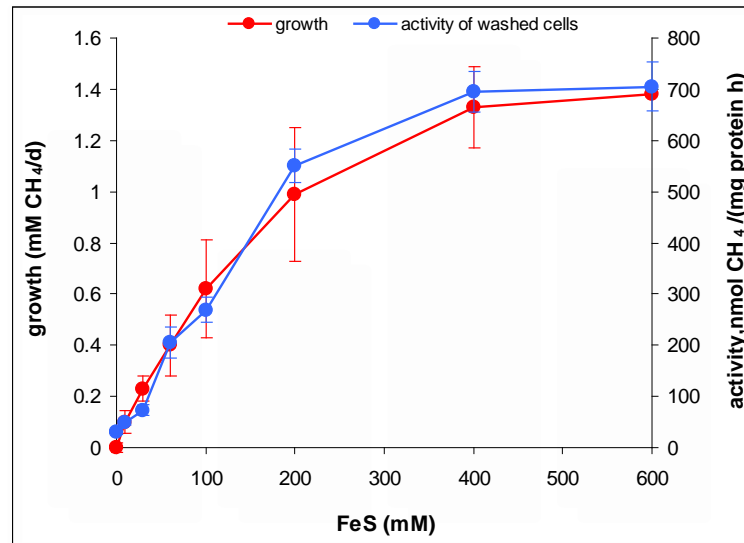


Fig.3

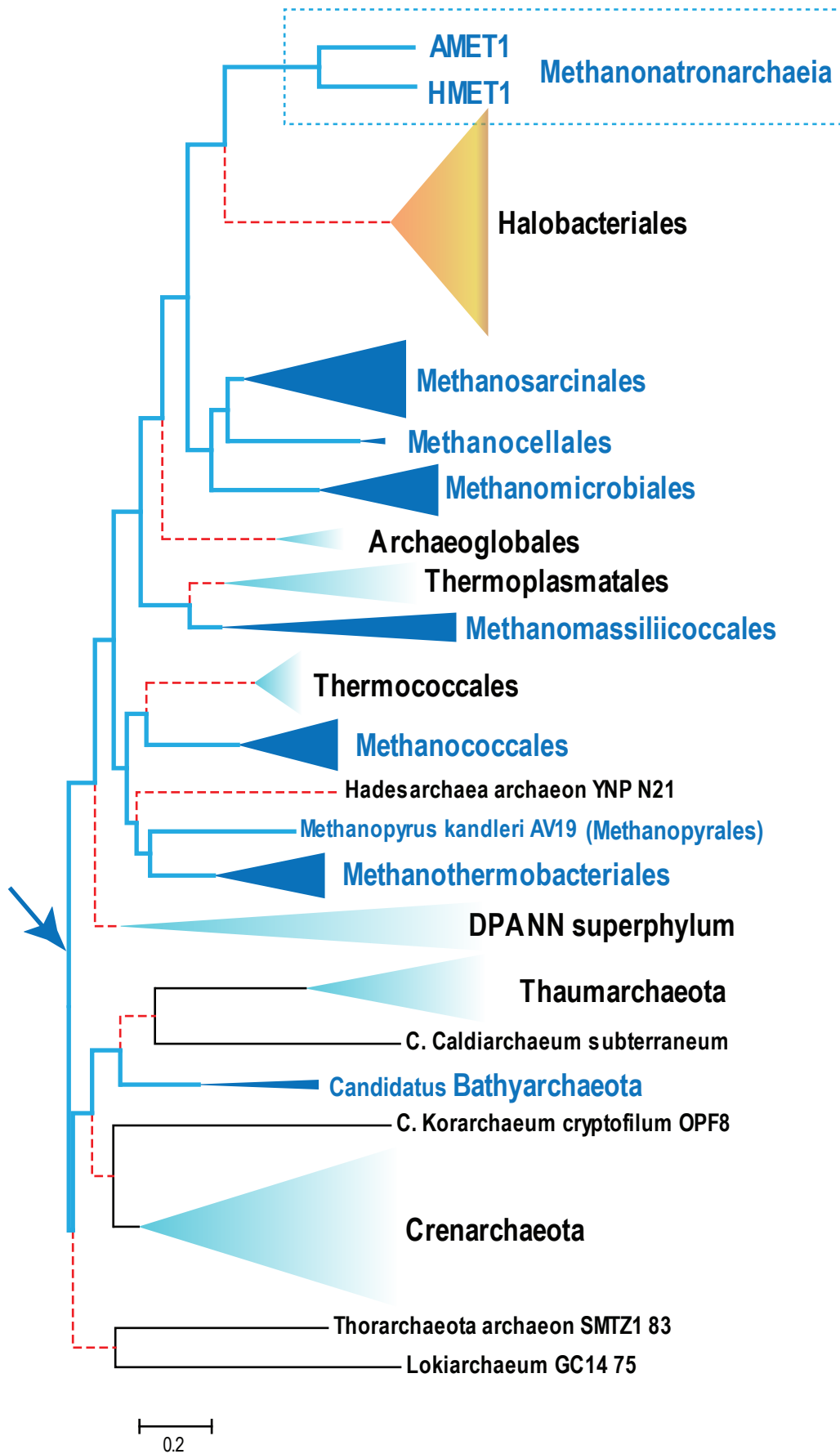
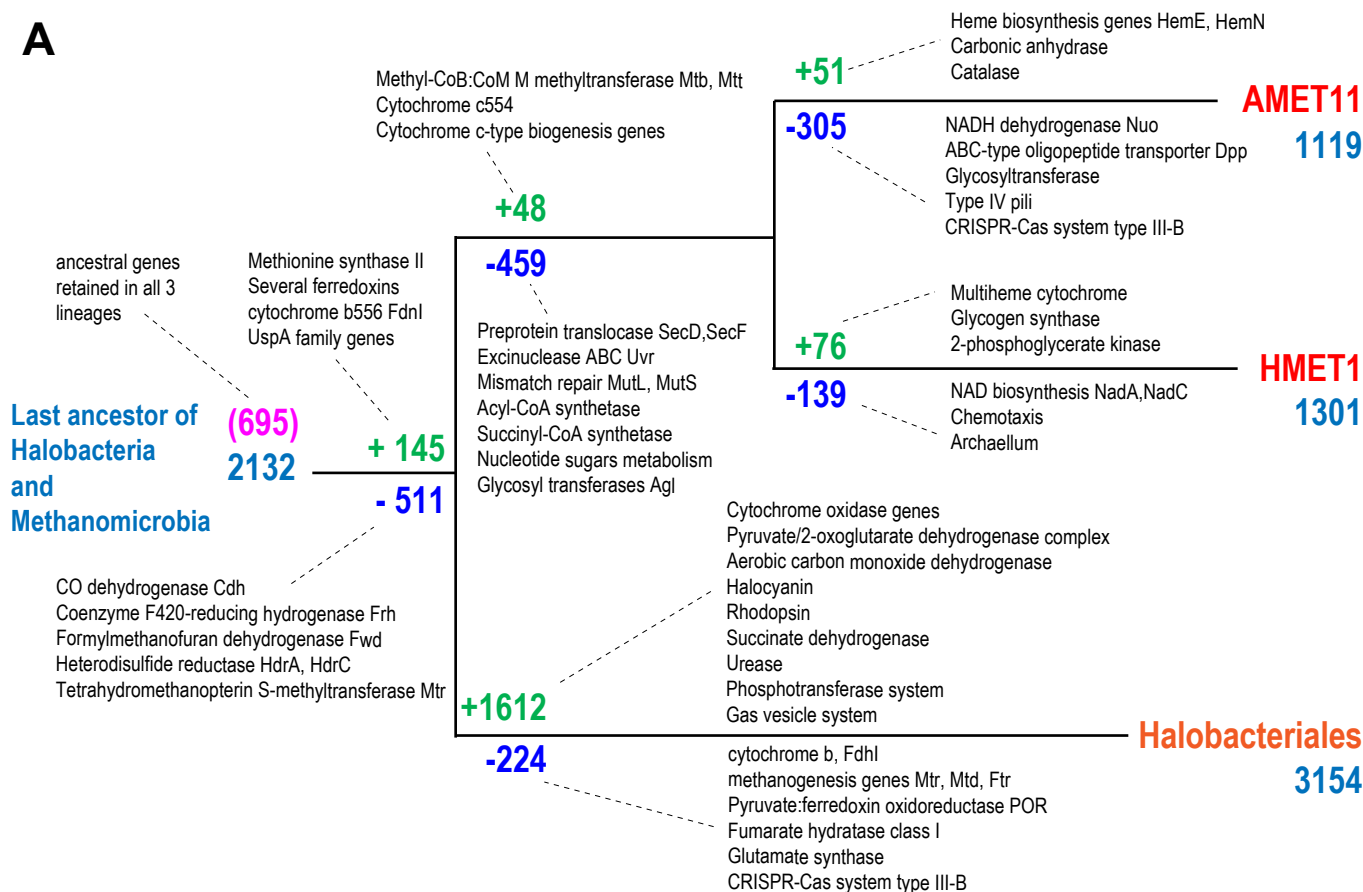
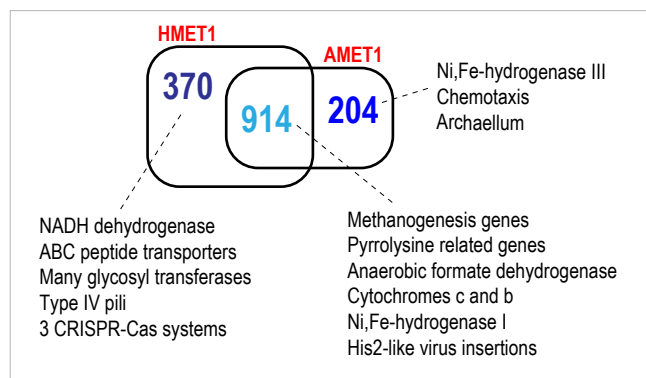


Fig. 4

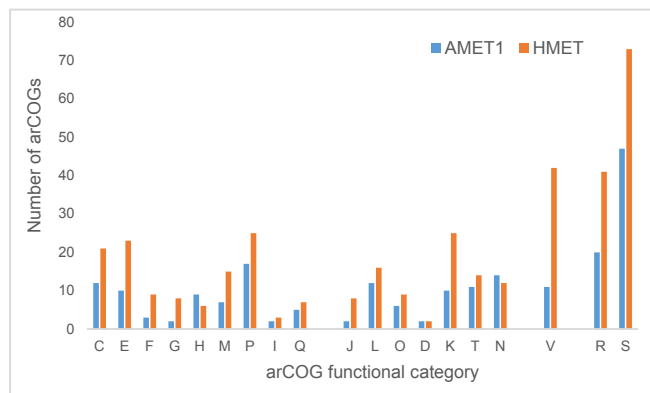
A



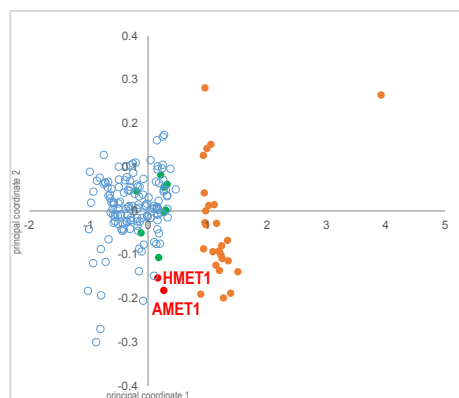
B



C



D



E

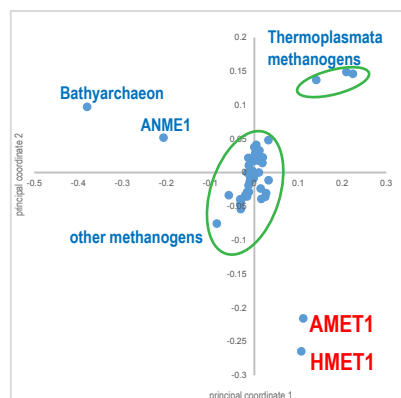


Fig. 5

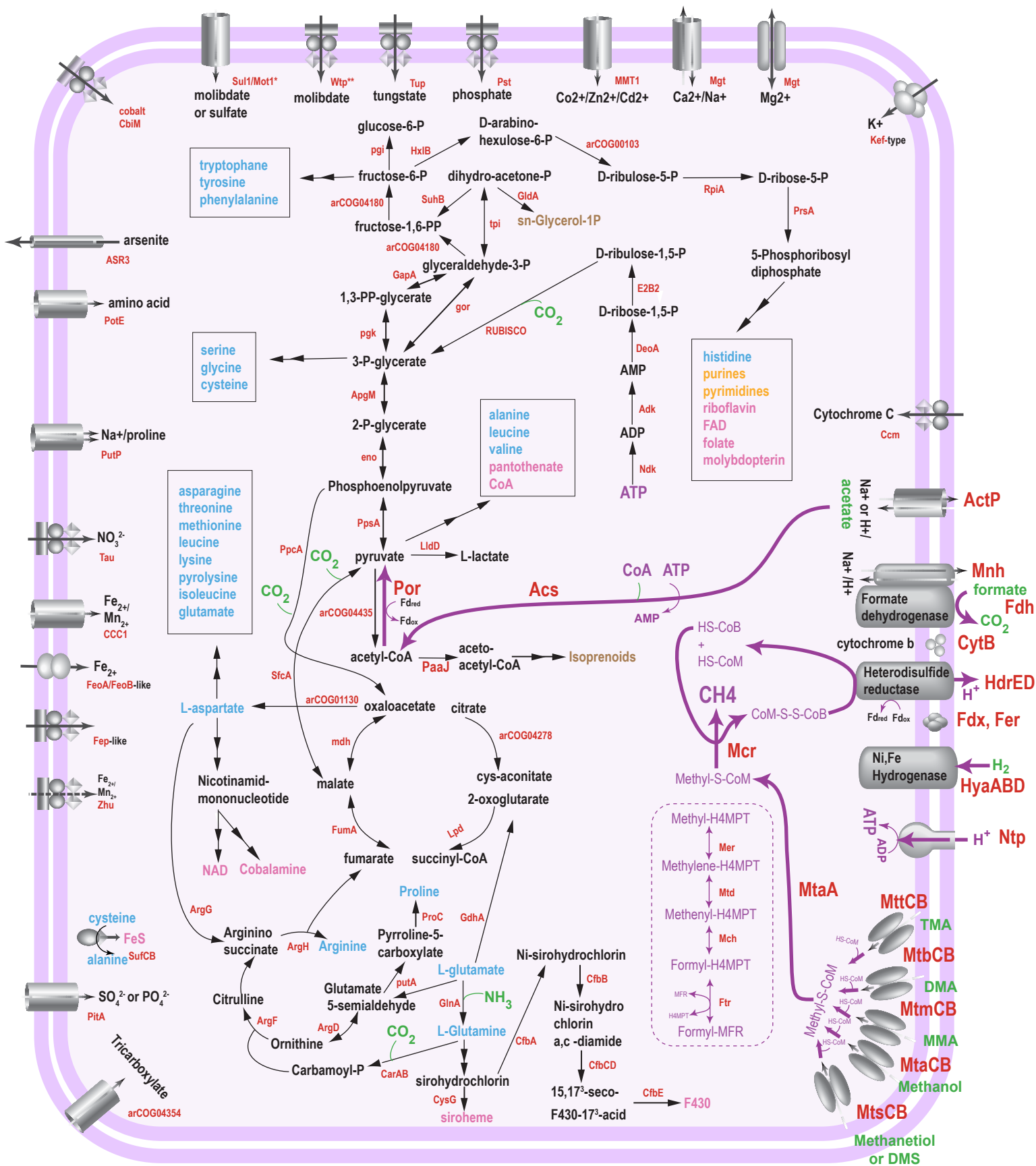


Fig. 6